AMERICAN SOCIETY FOR MICROBIOLOGY | Applied and Environmental Microbiology®

BIOTECHNOLOGY

Terminal Olefin Profiles and Phylogenetic Analyses of Olefin Synthases of Diverse Cyanobacterial Species

Tao Zhu,[a] Thibault Scalvenzi,[b] Nathalie Sassoon,[b] Xuefeng Lu,[a,c] Muriel Gugger[b]

[a]Key Laboratory of Biofuels, Shandong Provincial Key Laboratory of Synthetic Biology, Qingdao Institute of Bioenergy and Bioprocess Technology, Chinese Academy of Sciences, Qingdao, China

[b]Collection of Cyanobacteria, Department of Microbiology, The Institut Pasteur, Paris, France

[c]Laboratory for Marine Biology and Biotechnology, Qingdao National Laboratory for Marine Science and Technology, Qingdao, China

ABSTRACT Cyanobacteria can synthesize alkanes and alkenes, which are considered to be infrastructure-compatible biofuels. In terms of physiological function, cyanobacterial hydrocarbons are thought to be essential for membrane flexibility for cell division, size, and growth. The genetic basis for the biosynthesis of terminal olefins (1-alkenes) is a modular type I polyketide synthase (PKS) termed olefin synthase (Ols). The modular architectures of Ols and structural characteristics of alkenes have been investigated only in a few species of the small percentage (approximately 10%) of cyanobacteria that harbor putative Ols pathways. In this study, investigations of the domains, modular architectures, and phylogenies of Ols in 28 cyanobacterial strains suggested distinctive pathway evolution. Structural feature analyses revealed 1-alkenes with three carbon chain lengths (C_{15}, C_{17}, and C_{19}). In addition, the total cellular fatty acid profile revealed the diversity of the carbon chain lengths, while the fatty acid feeding assay indicated substrate carbon chain length specificity of cyanobacterial Ols enzymes. Finally, *in silico* analyses suggested that the N terminus of the modular Ols enzyme exhibited characteristics typical of a fatty acyl-adenylate ligase (FAAL), suggesting a mechanism of fatty acid activation via the formation of acyl-adenylates. Our results shed new light on the diversity of cyanobacterial terminal olefins and a mechanism for substrate activation in the biosynthesis of these olefins.

IMPORTANCE Cyanobacterial terminal olefins are hydrocarbons with promising applications as advanced biofuels. Despite the basic understanding of the genetic basis of olefin biosynthesis, the structural diversity and phylogeny of the key modular olefin synthase (Ols) have been poorly explored. An overview of the chemical structural traits of terminal olefins in cyanobacteria is provided in this study. In addition, we demonstrated by *in vivo* fatty acid feeding assays that cyanobacterial Ols enzymes might exhibit substrate carbon chain length specificity. Furthermore, by performing bioinformatic analyses, we observed that the substrate activation domain of Ols exhibited features typical of a fatty acyl-adenylate ligase (FAAL), which activates fatty acids by converting them to fatty acyl-adenylates. Our results provide further insight into the chemical structures of terminal olefins and further elucidate the mechanism of substrate activation for terminal olefin biosynthesis in cyanobacteria.

KEYWORDS cyanobacteria, biofuels, terminal olefins, olefin synthase, fatty acyl-adenylate ligase

Hydrocarbon biofuels are renewable candidates that have shown great potential to supplement the limited availability of petroleum in the future (1, 2). Two hydrocarbon biosynthetic pathways have been identified in cyanobacteria. The first pathway

Received 19 February 2018 **Accepted** 25 April 2018

Accepted manuscript posted online 4 May 2018

Citation Zhu T, Scalvenzi T, Sassoon N, Lu X, Gugger M. 2018. Terminal olefin profiles and phylogenetic analyses of olefin synthases of diverse cyanobacterial species. Appl Environ Microbiol 84:e00425-18. https://doi.org/10.1128/AEM.00425-18.

Editor Hideaki Nojiri, University of Tokyo

Address correspondence to Xuefeng Lu, lvxf@qibebt.ac.cn, or Muriel Gugger, muriel.gugger@pasteur.fr.

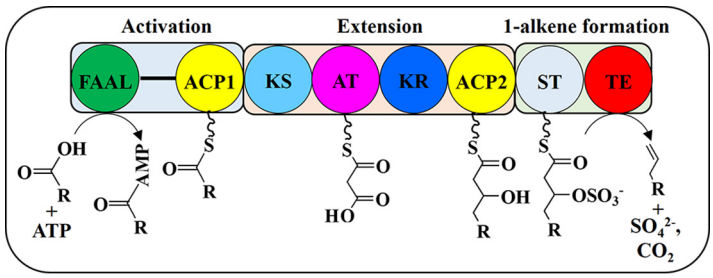

FIG 1 Schematic diagram of the cyanobacterial Ols pathway. The ATP-consuming fatty acyl-adenylate ligase (FAAL) domain first activates free fatty acid by conversion to fatty acyl-adenylate (fatty acyl-AMP), which is then transferred to the extension modules of the polyketide synthase (PKS)-like Ols enzyme. The acyl substrate is elongated by two carbons from malonyl-coenzyme A (CoA) via ketosynthase (KS) and acyl transferase (AT), which is followed by reduction to the β-hydroxyacid by a ketoreductase (KR). The 1-alkene formation modules are composed of sulfotransferase (ST) and thioesterase (TE), of which ST activates the β-hydroxy group via sulfonation, and TE activates subsequent dehydration and decarboxylation reactions to form the terminal alkene.

synthesizes alkanes and alkenes from a fatty acyl-acyl carrier protein (ACP), an intermediate of fatty acid metabolism, using two key enzymes, an acyl-ACP reductase (Aar) and an aldehyde-deformylating oxygenase (Ado) (1, 3). This pathway is present in a majority of cyanobacterial genomes (4) and produces a variety of hydrocarbons (saturated alkanes, branched alkanes, and unsaturated alkenes with internal double bonds) (2). In a small proportion of cyanobacteria, another pathway, termed the terminal olefin synthase (Ols) pathway, is present to synthesize alkenes (5, 6). This pathway is composed of a large type I polyketide synthase (PKS) with a modular organization and catalyzes the formation of a 1-alkene from fatty acids via an elongation-decarboxylation mechanism (6) (Fig. 1). The two hydrocarbon-producing pathways (Ado-Aar pathway and Ols pathway) seem to be mutually exclusive in all publicly available cyanobacterial genomes, which is thought to be associated with an unknown selective pressure (5). A previous comparison of average hydrocarbon yields from these pathways revealed that strains harboring the Ols pathway produced significantly more hydrocarbons than strains possessing the Ado-Aar pathway (0.17% and 0.07% of dry biomass, respectively) (5). Thus, it is of great biotechnological significance to further explore cyanobacterial strains with high terminal olefin production.

The N terminus of Ols, comprising a loading domain and an acyl carrier protein (ACP1), is predicted to function as the substrate activation module. The extension and C terminus of Ols, comprising ketosynthase (KS), acyltransferase (AT), ketoreductase (KR), ACP2, sulfotransferase (ST), and thioesterase (TE), are responsible for polyketide elongation and the formation of 1-alkene (5, 6) (Fig. 1). In most Ols-containing cyanobacteria, the Ols pathway is present in one open reading frame (ORF) (5). Deletions of the Ols pathway or of the N terminus of Ols have been shown to be essential in hydrocarbon synthesis only in the model microorganism *Synechococcus* sp. strain PCC 7002 (6). A feeding assay of *Synechococcus* sp. PCC 7002 revealed the substrate chain length specificity of cyanobacterial Ols (6). Further studies on strains containing this pathway are required to determine whether this specificity is a general property of Ols.

The subcellular location of hydrocarbons is within the lipid bilayers of thylakoid and cytoplasmic membranes, and the biological function of hydrocarbons is thought to be associated with the membrane flexibility required for optimal cell division, size, and growth (7). Free fatty acids in cyanobacteria are thought to be released from complex membrane lipids (8). The mechanism by which fatty acids are activated in the biosynthesis of terminal olefins has not been investigated thus far. A previous study elucidated a mode of fatty acid activation involving the formation of acyl-ACP thioesters catalyzed by acyl-ACP synthetase (Aas) in cyanobacteria (8). However, the Ols pathway exhibits a domain architecture similar to that of the fatty acyl-adenylate ligase (FAAL)-PKS pathway in *Mycobacterium tuberculosis* (Fig. 1). Intriguingly, the latter pathway is involved with the formation of fatty acyl-adenylate (AMP) in a mode of fatty acid

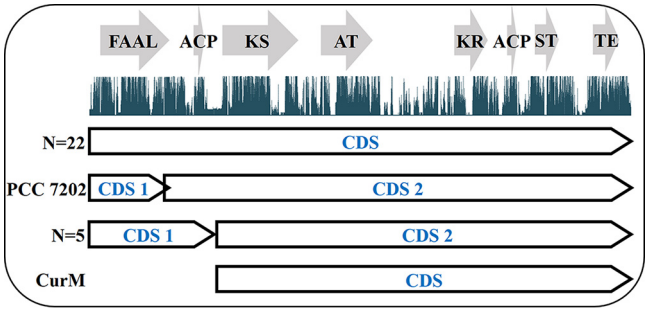

FIG 2 Deduced amino acid sequence alignment of 28 Ols enzymes and CurM. Twenty-two strains harbor complete FAAL-ACP1-KS-AT-KR-ACP2-ST-TE modules. *Cyanobacterium stanieri* PCC 7202 and five strains harbor the complete modules divided into two contiguous amino acid sequences (CDS 1 and 2). In *Cyanobacterium stanieri* PCC 7202, CDS 1 contains the FAAL domain, and CDS 2 contains the remaining domains. In the five remaining strains, CDS 1 contains the FAAL and ACP1 domains, and CDS 2 contains the remaining domains. The CurM domain of the curacin A biosynthetic pathway was the closest relative. List of the 22 strains with complete module: *Chroococcidiopsis* sp. PCC 6712, *Cyanobacterium* sp. IPPAS B-1200, cyanobacterium ESFC-1, *Cyanothece* sp. PCC 7424 and PCC 7822, *Geminocystis herdmanii* PCC 6308, *Geminocystis* sp. NIES-3709, *Leptolyngbya* sp. PCC 7376, *Limnothrix rosea* IAM M-220, *Myxosarcina* sp. GI1, *Pleurocapsa* sp. PCC 7319 and PCC 7327, *Stanieria cyanosphaera* PCC 7437, *Stanieria* sp. NIES-3757, *Synechococcus* sp. PCC 7002, PCC 7003, PCC 7117, PCC 8807, PCC 73109, NKBG042902, and NKBG15041c, and *Xenococcus* sp. PCC 7305. List of the 5 strains with CDS 1 and 2: *Leptolyngbya* sp. PCC 6406, *Moorea bouillonii* PNG5-198, and *Moorea producens* 3L, JHB, and PAL-8-15-08-1. The NCBI accession numbers of the Ols enzymes are indicated in Table 1.

activation (9, 10). This observation has led to the reclassification of universal fatty acid degradation enzymes, initially annotated as FadD, into FAALs and fatty acyl-coenzyme A ligases (FACLs) (9, 10). Evolved from FACLs, FAALs differ from FACLs by a typical insertion motif (10–12). Due to the similar architectures of FAAL-PKS modules in mycobacteria and cyanobacteria, these findings might be helpful for understanding the corresponding substrate activation of Ols in cyanobacteria.

Our study aimed to explore the terminal olefins of diverse cyanobacteria with a focus on the carbon chain length of the products. To study the substrate specificity of Ols, the capacities of representative strains to use exogenous fatty acids for terminal olefin biosynthesis were compared. Finally, we attempted to provide a basic understanding of substrate activation in cyanobacterial terminal olefin synthesis via bioinformatics analyses.

RESULTS

Domains and modular architectures of Ols in cyanobacteria. To investigate the domains and modular architectures of Ols, a genomic comparison of Ols pathways was conducted. Twenty-eight Ols candidates, including the Ols of eight strains with characterized terminal olefin compositions (6, 13), were selected. Using the CurM domain of the curacin A biosynthetic pathway of *Lyngbya majuscula* as a control, multiple sequence alignments were performed. We found that 22 strains contained the complete Ols enzyme in one ORF, while the remaining six strains harbored the modules in two contiguous coding sequences (CDS 1 and 2) (Fig. 2). In addition, the modular architecture of CDS 1 in *Cyanobacterium stanieri* PCC 7202 was different, with only an FAAL domain present instead of an FAAL domain and an ACP1 domain, both of which were present in the five other strains (Fig. 2). A close examination of the N termini of the Ols protein sequences revealed the presence of a shorter FAAL, with a 19-bp overlap with the adjacent PKS module, in *Cyanobacterium stanieri* PCC 7202 than in the other strains (see Fig. S1a and b in the supplemental material). The corresponding CDS 1 and CDS 2 nucleotide sequences, with 9- to 23-bp spacers, were found in the other five strains (Fig. S1a and c). Thus, the amino acid sequences of the Ols enzymes showed distinctive modular architectures in cyanobacteria.

Cyanobacteria are capable of producing three types of terminal olefins. After recognizing the novel features of the domains and modular architectures of Ols, 16

strains with morphotypes belonging to unicellular cyanobacteria, colonial unicellular cyanobacteria forming baeocytes, and filamentous cyanobacteria were selected further for an investigation of hydrocarbons (Table 1). Terminal olefins belonging to three classes of carbon chain length were produced by these strains (Fig. 3). The strains producing alkenes with double bonds at the first carbons, such as 1-pentadecene ($C_{15:1}$, Δ1) and 2-pentadecene ($C_{15:1}$, Δ2), were *Chroococcidiopsis* sp. strain PCC 6712 and *Xenococcus* sp. strain PCC 7305 (Fig. 3; see also Fig. S2a and b). Seven of the cyanobacteria produced 1-heptadecene ($C_{17:1}$, Δ1) (Fig. 3). *Geminocystis herdmanii* strain PCC 6308 and *Cyanobacterium stanieri* PCC 7202 synthesized hydrocarbons with novel structural features (2-, 3-, and 4-heptadecene), while *Pleurocapsa* sp. strain PCC 7319 produced the most diverse hydrocarbons, such as 1-, 2-, 3-, 4-, and 5-heptadecene, as confirmed by dimethyl disulfide (DMDS) derivatization and subsequent gas chromatography-mass spectrometry (GC-MS) analysis (Fig. 3; see also Fig. S3a, b, c, d, and e). Moreover, seven strains were found to produce C_{19} terminal olefins (Fig. 3; Fig. S2c). Mass spectra of $C_{19:2}$ hydrocarbons were observed for all C_{19} terminal olefin-producing strains (Fig. 3; Fig. S2d). In terms of hydrocarbon yield, *Cyanothece* sp. strain PCC 7424 exhibited a production of ~0.44% of the lyophilized biomass (Fig. 3), which is approximately 1.7 times the highest terminal olefin production reported in a native cyanobacterial strain (0.26% of dry biomass) (5).

Phylogeny of olefin synthase in cyanobacteria. A phylogenetic tree of the Ols pathway and the CurM domain of the curacin A biosynthetic pathway was generated using the maximum likelihood method (Fig. 4). The five strains harboring separate FAAL-ACP1 modules and CurM domains grouped into one cluster, which was separated from the cluster of strains harboring complete Ols pathways and *Cyanobacterium stanieri* PCC 7202. Taken together, the terminal olefin composition data (Fig. 3) and the data regarding Ols domain architecture and phylogeny reveal that the strains with distinctive Ols pathways can produce both common hydrocarbons (such as 1-nonadecene in *Leptolyngbya* sp. strain PCC 6406) and novel alkenes (such as 1-, 2-, 3,- and 4-heptadecene in *Cyanobacterium stanieri* PCC 7202) (Fig. 3). Meanwhile, strains with complete Ols pathways are also capable of producing common and novel hydrocarbons (Fig. 3). Thus, the different domain architectures and/or phylogenies of Ols might not necessarily correspond to a phenotype of novel alkene structure. In terms of alkene chain length, the clustering characteristics of the hydrocarbon profile and Ols phylogeny appeared to exhibit some correlation, but additional samples are required to draw a conclusion.

Total fatty acid composition of the 16 terminal olefin-producing cyanobacterial strains. The precursors of terminal olefin synthesis are derivatives of fatty acids. Complex membrane lipids contribute to the dynamic pool of free fatty acids in cyanobacteria. Thus, fatty acid compositions in the total lipid fractions of the 16 strains were investigated using a methyl esterification method.

In C_{15} terminal olefin-producing strains (*Chroococcidiopsis* sp. PCC 6712 and *Xenococcus* sp. PCC 7305), C_{14} fatty acids were synthesized, although C_{16} fatty acids were in high abundance (Fig. 5). Interestingly, the concentration of C_{14} fatty acids in *Chroococcidiopsis* sp. PCC 6712 was 38-fold higher than in *Xenococcus* sp. PCC 7305 (Fig. 5), yet the latter strain produced more C_{15} hydrocarbons than the former (Fig. 3). In most of the C_{17} terminal olefin-producing strains, large amounts of C_{16} fatty acids and small amounts of C_{18} fatty acids were produced (Fig. 5). In addition, C_{14} fatty acids were also found in these strains, with a proportion of up to ~36% of the total fatty acids in *Geminocystis herdmanii* PCC 6308 and ~20% in *Cyanobacterium stanieri* PCC 7202 (Fig. 5). In C_{19} terminal olefin-producing strains, the key components of the total fatty acids were C_{18} and C_{16} fatty acids, but no C_{14} fatty acids were identified (Fig. 5). Thus, the total fatty acid profiles of cyanobacteria are closely associated with olefin carbon chain length and might be one of the factors influencing hydrocarbon chain length specificity. However, the abundances of these fatty acids seem to not be directly associated with the production of hydrocarbons.

TABLE 1 Cyanobacterial strains and GenBank accession numbers of Ols, FAALs, and FACLs used for bioinformatic analyses in this study

Subsection (order), morphotype	Organism	Protein	NCBI accession no.
Cyanobacteria			
Subsection I (*Chroococcales*), unicellular	*Cyanobacterium stanieri* PCC 7202	Ols	AFZ46265.1
		FAAL	AFZ46264.1
	Cyanobacterium sp. IPPAS B-1200	Ols	WP_071891741.1
	Cyanothece sp. PCC 7424	Ols	WP_012599249.1
	Cyanothece sp. PCC 7822	Ols	WP_013321938.1
	Geminocystis herdmanii PCC 6308	Ols	WP_017293948.1
	Geminocystis sp. NIES-3709	Ols	WP_060833030.1
	Synechococcus sp. NKBG042902	Ols	WP_030006270.1
	Synechococcus sp. NKBG15041c	Ols	WP_024544961.1
	Synechococcus sp. PCC 7002	Ols	WP_012306795.1
	Synechococcus sp. PCC 7003	Ols	WP_065713741.1
	Synechococcus sp. PCC 7117	Ols	WP_065710482.1
	Synechococcus sp. PCC 73109	Ols	WP_062433466.1
	Synechococcus sp. PCC 8807	Ols	WP_065716260.1
Subsection II (*Pleurocapsales*), baeocytous	*Chroococcidiopsis* sp. PCC 6712	Ols	2505786305[a]
	Myxosarcina sp. GI1	Ols	WP_072013783.1
	Pleurocapsa sp. PCC 7319	Ols	WP_019509160.1
	Pleurocapsa sp. PCC 7327	Ols	AFY79044.1
	Stanieria cyanosphaera PCC 7437	Ols	AFZ37598.1
	Stanieria sp. NIES-3757	Ols	BAU65096.1
	Xenococcus sp. PCC 7305	Ols	WP_006509673.1
Subsection III (*Oscillatoriales*), filamentous	Cyanobacterial strain ESFC-1	Ols	WP_018399776.1
	Leptolyngbya sp. PCC 6406	Ols	WP_008319224.1
		FAAL	WP_008319225.1
	Leptolyngbya sp. PCC 7376	Ols	WP_015132477.1
	Limnothrix rosea IAM M-220	Ols	OKH15930.1
	Lyngbya majuscula	CurA	AEE88277.1
	Moorea bouillonii PNG5-198	Ols	AHH34186.1
		FAAL	AHH34187.1
	Moorea producens 3L	Ols	EGJ35088.1
		FAAL	EGJ35087.1
	Moorea producens JHB	Ols	AHH34189.1
		FAAL	AHH34188.1
	Moorea producens PAL-8-15-08-1	Ols	AOW99348.1
		FAAL	AOW99347.1
Actinobacteria and *Proteobacteria*			
Mycobacteria	*Mycobacterium tuberculosis*	FACL1	NP_216266.1
	Mycobacterium tuberculosis	FACL2	NP_214784.1
	Mycobacterium tuberculosis	FACL3	NP_218078.1
	Mycobacterium tuberculosis	FACL4	NP_214728.1
	Mycobacterium tuberculosis	FACL5	NP_214680.1
	Mycobacterium tuberculosis	FACL6	NP_215722.1
	Mycobacterium tuberculosis	FACL7	NP_214633.1
	Mycobacterium tuberculosis	FACL8	NP_215065.1
	Mycobacterium tuberculosis	FACL12	NP_215943.1
	Mycobacterium tuberculosis	FACL13	NP_217605.1
	Mycobacterium tuberculosis	FACL14	NP_215574.1
	Mycobacterium tuberculosis	FACL15	NP_216703.1
	Mycobacterium tuberculosis	FACL17	NP_218023.1
	Mycobacterium tuberculosis	FACL19	YP_177983.1
	Mycobacterium tuberculosis	FACL22	NP_217464.1
	Mycobacterium tuberculosis	FACL34	YP_177686.1
	Mycobacterium tuberculosis	FACL35	NP_217021.1
	Mycobacterium tuberculosis	FACL36	NP_215709.1
	Mycobacterium tuberculosis	FAAL10	NP_214613.1
	Mycobacterium tuberculosis	FAAL21	NP_215701.1
	Mycobacterium tuberculosis	FAAL23	NP_218343.1
	Mycobacterium tuberculosis	FAAL24	NP_216045.1
	Mycobacterium tuberculosis	FAAL25	NP_216037.1
	Mycobacterium tuberculosis	FAAL26	NP_217446.2
	Mycobacterium tuberculosis	FAAL28	NP_217457.1
	Mycobacterium tuberculosis	FAAL29	NP_217466.3
	Mycobacterium tuberculosis	FAAL30	NP_214918.1
	Mycobacterium tuberculosis	FAAL31	NP_216441.2
	Mycobacterium tuberculosis	FAAL32	NP_218318.1
	Mycobacterium tuberculosis	FAAL33	NP_215861.1
Enterobacterales	*Escherichia coli* CFT073	FAAL	NP_755583.1
Legionellales	*Legionella pneumophila*	FAAL	YP_096241.1
Myxococcales	*Myxococcus xanthus* DK 1622	FAAL	YP_634753.1

[a]JGI IME ID number of Ols from *Chroococcidiopsis* sp. PCC 6712.

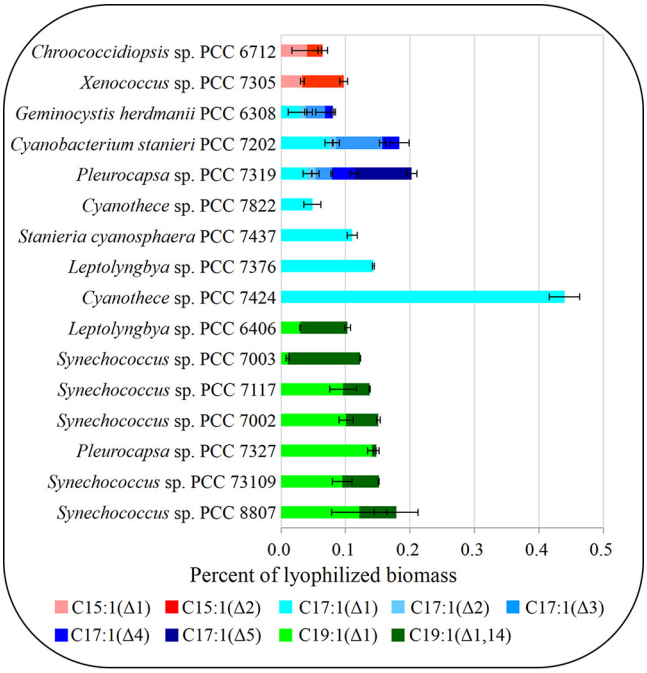

FIG 3 Hydrocarbon profile of 16 cyanobacterial strains capable of producing terminal olefins. Terminal olefins with chain lengths of C_{15}, C_{17}, and C_{19} were identified; the number after the colon indicates the number of double bonds, while the number after the triangle indicates the position of the double bond. The 1,14-nonadecadienes ($C_{19:2}$, Δ1,14) were identified on the basis of their mass spectra. Data shown are the means from three biological replicates, and the error bars represent standard deviations.

Cyanobacteria exhibit fatty acid chain length specificity for terminal olefin production. To further investigate the relationship between hydrocarbon chain length and Ols enzymes, we carried out an odd-chain fatty acid feeding assay. Three representative strains (*Chroococcidiopsis* sp. PCC 6712, *Stanieria cyanosphaera* PCC 7437, and *Synechococcus* sp. PCC 7002), producing C_{15}, C_{17}, and C_{19} terminal olefins, were selected for the feeding assay. A supplementation with nonadecanoic acid ($C_{19:0}$) did not alter the hydrocarbon compositions of the three strains (Fig. 6). A supplementation with heptadecanoic acid ($C_{17:0}$) led to the production of 1-octadecene ($C_{18:1}$, Δ1) by *Synechococcus* sp. PCC 7002 and *Stanieria cyanosphaera* PCC 7437 (Fig. 6). A supplementation with pentadecanoic acid ($C_{15:0}$) prompted the synthesis of only 1-octadecene by *Synechococcus* sp. PCC 7002, while this fatty acid enabled the production of both 1-octadecene and 1-hexadecene ($C_{16:1}$, Δ1) by *Stanieria cyanosphaera* PCC 7437 (Fig. 6). The Ols of *Chroococcidiopsis* sp. PCC 6712 appeared to be unable to use exogenous fatty acids with chain lengths of C_{15}, C_{17}, and C_{19} to produce C_{16} hydrocarbons or any other hydrocarbons under our experimental conditions. The results confirmed that cyanobacterial Ols pathways exhibit fatty acid chain length specificity for alkene biosynthesis, thus revealing increased diversity in the Ols pathways of cyanobacteria.

The loading domain of Ols is a typical fatty acyl-AMP ligase. *Stanieria cyanosphaera* PCC 7437, which produced $C_{17:1}$ hydrocarbons, is therefore capable of using $C_{15:0}$ acids for both $C_{16:1}$ and $C_{18:1}$ production. The exogenous fatty acids must be activated and incorporated into the Ols pathway. We investigated the possible mechanisms of fatty acid activation of Ols. The FAAL-PKS architecture of *Mycobacterium tuberculosis* (9) is similar to that of the cyanobacterial Ols pathway. The sequences of the FACLs and FAALs of *Mycobacterium tuberculosis*, the FAALs of *Escherichia coli*, *Legionella pneumophila*, and *Myxococcus xanthus*, and the FAALs of cyanobacteria were compared (Fig. 7). The cyanobacterial loading domain of Ols exhibited characteristics typical of FAAL enzymes, with similar signature insertion motifs. A maximum likelihood

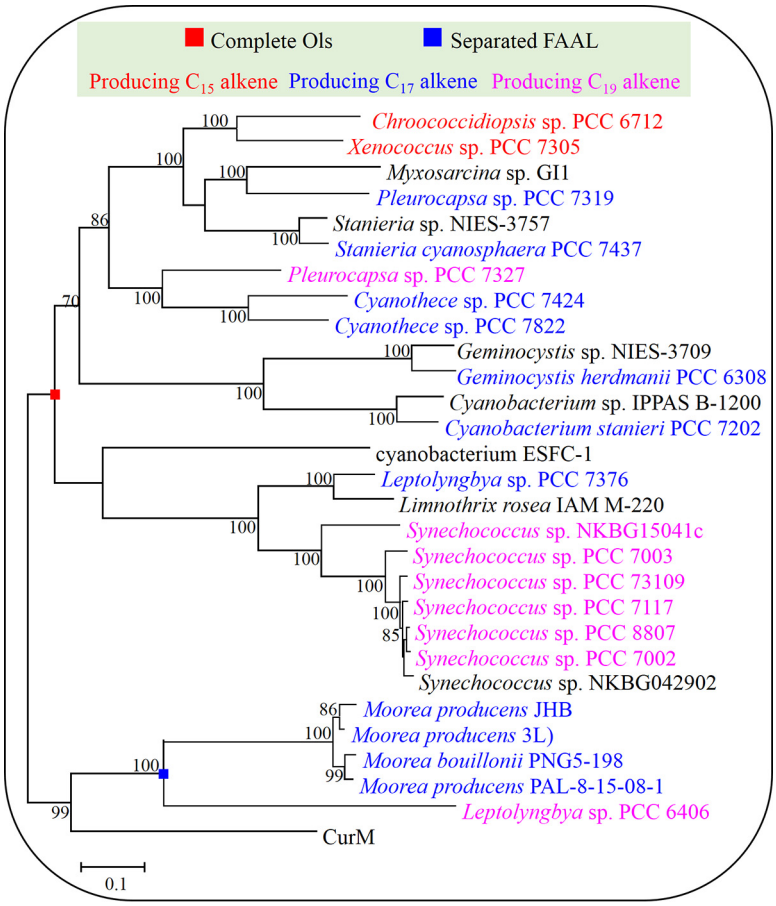

FIG 4 Phylogeny of the Ols pathways of 28 cyanobacterial strains. The tree was generated by the maximum likelihood method. CurM was used as the outgroup. Bootstrap values >70%, expressed as percentages of 1,000 replications, are shown near the nodes. The clade labeled with a red square has a complete Ols, and in the clade labeled with a blue square, the FAAL-ACP1 is separated from the PKS. Strain *Cyanobacterium stanieri* PCC 7202, possessing distinctive Ols modular architecture, was grouped into the clade with the red square. Red, blue, and pink strain names indicate Ols pathways corresponding to alkenes with chain lengths of C_{15}, C_{17}, and C_{19}, respectively. The NCBI accession numbers of Ols and CurM are indicated in Table 1.

phylogeny based on these FAAL and FACL sequences confirmed that cyanobacterial and mycobacterial FAALs are clustered and are separated from mycobacterial FACLs (see Fig. S4). Thus, the loading domain of Ols is that of a typical FAAL, and the activated form of the substrate in cyanobacterial terminal olefin synthesis could be fatty acyl-AMP as well.

DISCUSSION

Hydrocarbons are widespread in cyanobacteria, and the cyanobacteria sequenced to date possess one of the two separate hydrocarbon-producing pathways (2, 5). In this study, five of the 28 Ols-containing cyanobacterial strains were found to harbor separate FAAL-ACP1 modules, and together with the peculiar Ols domain architecture of *Cyanobacterium stanieri* PCC 7202, these results reveal the genetic diversity of Ols enzymes in cyanobacteria (Fig. 2). The independent clades in the phylogeny of Ols revealed a possible distinct evolutionary history of these two groups of Ols (Fig. 4).

The physiological roles of cyanobacterial hydrocarbons, which are involved in introducing flexibility into membranes and are required for optimal cell division, size, and growth, have been previously reported (7). Though strains harboring the Ols pathway have been proposed to account for a small proportion (approximately 10%) of the cyanobacterial population (4, 5), hydrocarbon production in such strains is recog-

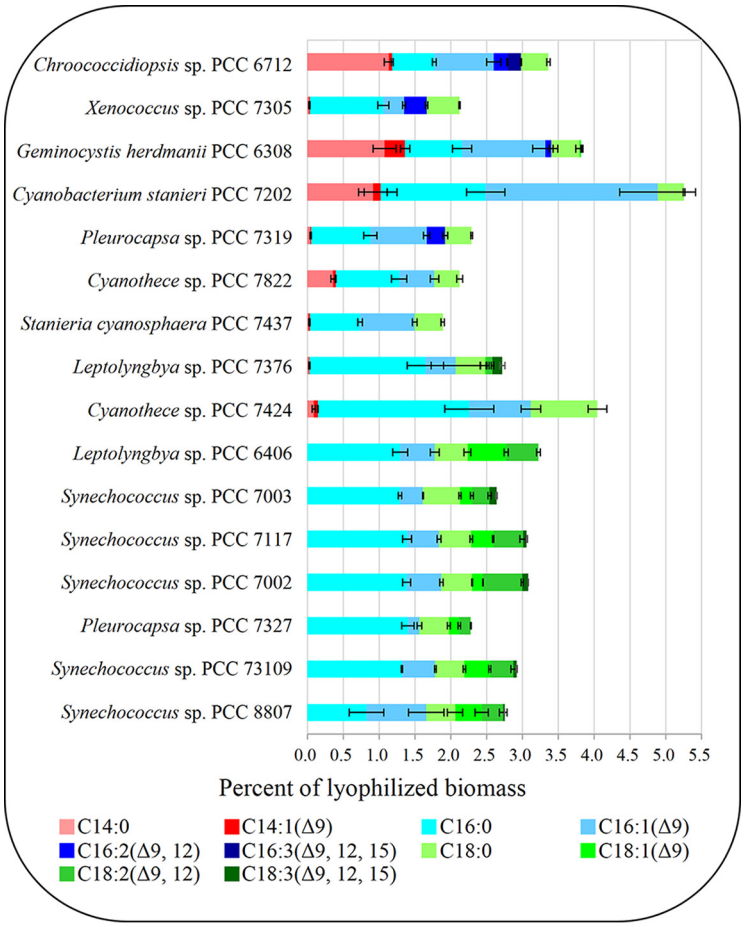

FIG 5 Total fatty acid composition of the 16 terminal olefin-producing cyanobacterial strains. Fatty acids with chain lengths of C_{14}, C_{16}, and C_{18} were identified; the number after the colon indicates the number of double bonds, while the number after the triangle indicates the position of the double bond. Data are presented as the weight percentages of corresponding fatty acids. Data shown are the means from three biological replicates, and the error bars represent standard deviations.

nized to be higher than that in strains containing the Ado/Aar pathway (2, 5). In this study, we found that *Cyanothece* sp. PCC 7424 exhibits higher hydrocarbon production (0.44% of the lyophilized biomass) than the reported wild-type-terminal olefin-producing cyanobacteria (Fig. 3). To obtain an overview of the terminal olefin profiles of cyanobacteria, we collected a majority of strains that, based on genome analyses, contained Ols and conducted rigorous structural identification of the hydrocarbons produced by these strains. The novel cyanobacterial 1-pentadecene ($C_{15:1}$, Δ1) and 2-pentadecene ($C_{15:1}$, Δ2) were identified in the cyanobacteria (Fig. 3). In addition, alkenes with internal double bonds, such as 2-petadecene and 2-, 3-, 4-, and 5-heptadecene, were identified in this study (Fig. 3). *Pleurocapsa* sp. PCC 7319 presented the highest diversity of Ols pathway hydrocarbons (Fig. 3). A desaturase gene (*desE*) was identified as being involved in the formation of the internal double bond of 1,14-nonadecadiene (14), providing a possible pathway for the generation of internal double bonds in hydrocarbons. Alternatively, isomerase could catalyze the formation of internal double bonds. The synthetic mechanisms for these types of hydrocarbons remain to be studied further. In addition, $C_{19:2}$ hydrocarbons were found in the seven C_{19} terminal olefin-producing strains explored in this study. On the basis of the retention times and mass spectra, we annotated the $C_{19:2}$ hydrocarbon as 1,14-nonadecadiene. Previous studies showed that the Ols of *Leptolyngbya* sp. PCC 6406 had a high GC content (64%) and suggested that the Ols of this strain was acquired via horizontal gene transfer (5). Our hydrocarbon analysis confirmed that

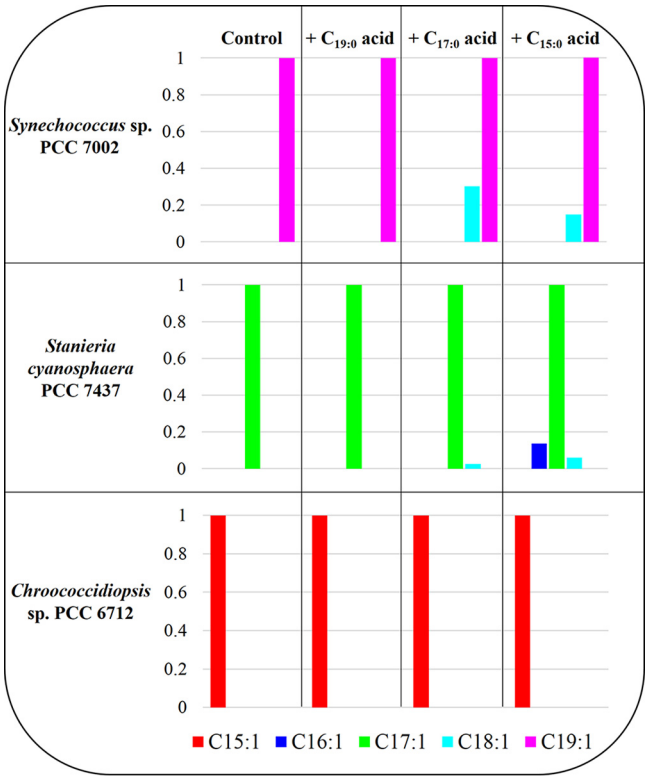

FIG 6 Odd-chain fatty acid feeding assay in representative cyanobacterial strains. The GC-MS signals were normalized to the peak areas of the native hydrocarbon of corresponding cyanobacterial strains. A supplementation of *Synechococcus* sp. PCC 7002 cultures with pentadecanoic acid ($C_{15:0}$) resulted in the synthesis of 1-octadecene ($C_{18:1}$, Δ1), but a supplementation of *Stanieria cyanosphaera* PCC 7437 cultures with pentadecanoic acid ($C_{15:0}$) led to the synthesis of both 1-hexadecene ($C_{16:1}$, Δ1) and 1-octadecene ($C_{18:1}$, Δ1). Fatty acid feeding assays were conducted with three replicates, and data shown are representative results from one replicate. Additionally, the experiments were also repeated using DMSO-dissolved and ethanol-dissolved fatty acids, and similar hydrocarbon profiles were obtained.

Leptolyngbya sp. PCC 6406 produces regular 1-nonadecene despite the evolutionary history of its Ols pathway.

Cellular fatty acid profiles can affect the carbon chain lengths of terminal olefins; C_{14} fatty acid was found only in C_{15} and C_{17} but not C_{19} terminal olefin-producing strains. In addition, large amounts of C_{16} fatty acid in C_{17} terminal olefin-producing strains and nearly similar levels of C_{16} and C_{18} fatty acids in C_{19} terminal olefin-producing strains further confirmed this effect on hydrocarbon chain length. However, no C_{15} terminal olefin was identified in C_{17} terminal olefin-producing strains, although large amounts of C_{14} fatty acid were found in some of these strains, indicating the substrate specificity of cyanobacterial Ols enzymes. Hence, this precursor specificity could be another factor that affects the carbon chain length of terminal olefins. The exogenous fatty acid feeding assay of strains with different alkene compositions provided the possibility to further explore the catalytic mechanism and substrate specificity of different Ols enzymes.

Little was known about the form of fatty acid activation in terminal olefin biosynthesis. The *in silico* analyses revealed that the loading domains of Ols have typical features of FAALs. To date, only FadD10 of *Mycobacterium tuberculosis*, which is missing the typical insertion sequences, has been shown to be a new FAAL (FAAL10) (12) (Fig. 7), and no FAAL enzyme shows FACL activity. Therefore, the FAAL domain of Ols can probably directly activate fatty acids to form fatty acyl-AMP by using ATP. Indeed, both FAALs and FACLs are capable of utilizing fatty acid pools and channeling these fatty acids toward different metabolic fates in mycobacteria (9–11). Even though the loading

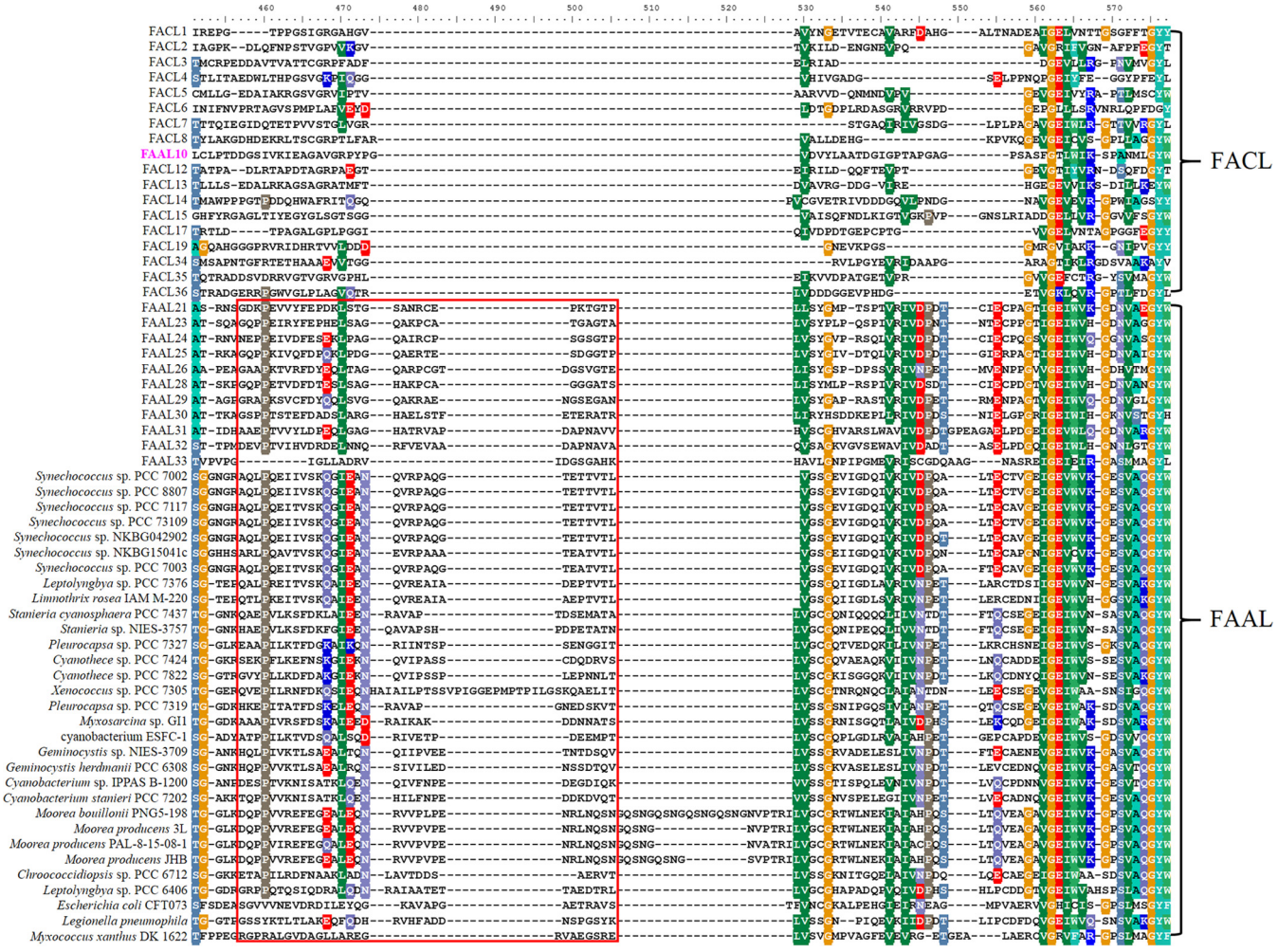

FIG 7 Demonstration of the typical insertion motif of cyanobacterial FAALs. Multiple sequence alignments of cyanobacterial FAALs, mycobacterial FAALs and FACLs, and FAALs of various proteobacteria were conducted by BioEdit software. The FAAL-specific insertion motif is highlighted with a red box. The NCBI accession numbers of Ols, FAALs, and FACLs are indicated in Table 1. FadD10 of *Mycobacterium tuberculosis* was reported to be a new type of FAAL even though it lacked the insertion sequence (12) (FAAL10 in bold pink).

domains of Ols enzymes are similar to those of FAALs, other mechanisms for fatty acid activation cannot be excluded. For example, the $C_{15:0}$ acid used for C_{18} alkene production in cyanobacteria favors the previously proposed mechanism of fatty acid activation catalyzed by Aas to form fatty acyl-ACP (6). A deletion mutant of the *aas* gene in *Synechococcus* sp. PCC 7002 is required to elucidate the mechanism, and further *in vitro* studies are needed for confirmation.

In summary, we obtained an overview of the terminal olefin profiles of cyanobacteria, identified several novel terminal olefins in cyanobacteria, and found that *Cyanothece* sp. PCC 7424 exhibits high terminal olefin production under our growth conditions. Meanwhile, analyses of modular architectures and phylogenies demonstrated two distinct evolutionary groups of Ols enzymes. Furthermore, by investigating the total fatty acid content and by a fatty acid feeding assay, we showed that the Ols enzymes of cyanobacteria have different substrate chain length specificities. Finally, we proposed that FAALs perform substrate activation in terminal olefin biosynthesis, most likely via the formation of fatty acyl-AMP. As fundamental research, we explored a new mode of fatty acid activation in cyanobacteria, and for biotechnological purposes, we demonstrated the substrate specificity of Ols, which is promising for the improvement of biofuel quality via FAAL domain substitutions.

MATERIALS AND METHODS

Chemical reagents. Eicosane and fatty acids were obtained from Sigma-Aldrich (USA). All other chemicals were obtained from either Merck (Germany) or Ameresco (USA).

Cyanobacterial strains, culture conditions, and biomass preparation. Three independent biological replicate cyanobacterial cultures (Table 1) were grown in 1.2 liters of BG11 medium (15) with 10 mM $NaHCO_3$ and 1% CO_2 (bubbled) under continuous light at 20 μmol photon $m^{-2} \cdot s^{-1}$ to obtain enough biomass for chemical analyses. The cells were centrifuged, washed twice with distilled water, and lyophilized.

Hydrocarbon extractions. Hydrocarbons were extracted as previously reported with minor modifications (16, 17). Fifty milligrams of lyophilized cyanobacterial biomass were resuspended in 10 ml of sterile deionized water and lysed by sonication. For strains of morphological subsections II and III (Table 1), the resuspended samples were homogenized by a 20-ml tissue homogenizer to assist the sonication. The lysate was extracted for 2 h at room temperature with 10 ml of chloroform-methanol (2:1 [vol/vol]). Prior to extraction, 30 μg of eicosane ($C_{20:0}$) was added to the cell lysate as an internal standard. The organic phase was separated following centrifugation (8,000 $\times$ *g*, 5 min) and evaporated to dryness under a nitrogen stream at 55°C. The residue containing the hydrocarbons was redissolved in 1 ml of *n*-hexane for further analysis.

Total lipid extraction and methyl esterification of fatty acids. Total lipids were extracted as previously reported with minor modifications (16–18). Twenty milligrams of lyophilized cyanobacterial biomass was transferred to a 15-ml tube and resuspended with 2 ml of sterile deionized water. For strains of morphological subsections II and III (Table 1), the resuspended samples were homogenized by a 5-ml tissue homogenizer to assist the lipid extraction. Prior to the extraction, 50 μg of nonadecanoic acid ($C_{19:0}$) was added to the cell suspension as the internal standard. The samples were extracted by adding 4 ml of chloroform-methanol (1:1 [vol/vol]), which was followed by homogenization of the cells with a vortex generator. The lower organic phase was separated by centrifugation (10,000 $\times$ *g*, 5 min), transferred to a 15-ml esterification tube, and evaporated to dryness under a nitrogen stream at 55°C. Then, 2 ml of 0.4 M KOH-methanol solution was added, and the mixture was incubated at 60°C for 1 h, allowing transesterification of lipid-bound fatty acids to the corresponding fatty acid methyl esters (FAMEs). Finally, 2 ml of *n*-hexane and 3 ml of 5 M NaCl were added and gently mixed, and after keeping the mixture at room temperature for 20 min, the FAMEs (upper hexane phases) were transferred to sample vials for further analysis.

Hydrocarbon and total fatty acid analyses. Hydrocarbons and total fatty acids were analyzed as previously reported with minor modifications (16, 17). One-microliter aliquots of the redissolved lipid samples were analyzed by GC-MS using an Agilent 7890A system equipped with an HP-INNOWax capillary column (30 m by 250 μm by 0.25 μm) and coupled to an Agilent 5975C MSD single-quadrupole mass spectrometer under electron ionization mode at 70 eV in a scan range of 50 to 500 *m/z*. Helium (constant flow rate of 1 ml $\cdot$ min^{-1}) was used as the carrier gas. The injection temperature was set to 250°C, and the injection volume was 1 μl, under splitless injection conditions. The following temperature program was applied: 40°C for 1 min, increased to 200°C at 5°C per min, increased to 240°C at 25°C per min, and held at 240°C for 15 min. For analyses of total fatty acids (the corresponding FAMEs), a similar temperature program was employed, with the exception that the initial temperature was 100°C. During GC-MS analysis, peak identification was conducted by comparing the retention times and fragmentation patterns with those of authentic standard compounds when available and by comparing with the values available in the mass spectral database of the National Institute of Standards and Technology, USA. The internal standard eicosane ($C_{20:0}$) was used to quantify the hydrocarbons, and methyl nonadecanoate (methyl ester of the internal standard nonadecanoic acid) was used to quantify FAMEs.

Derivatization with dimethyl disulfide. To determine the position of carbon-carbon double bonds of hydrocarbons, the DMDS derivatization method was employed with minor modifications (19). Extracted hydrocarbon samples (100 μl) were treated with 300 μl of DMDS (as a control, the hydrocarbon sample was treated with 300 μl of *n*-hexane) and 50 μl of iodine-diethyl ether solution (80 mg $\cdot$ ml^{-1} [m/vol]). The reaction mixture was allowed to stand for 6 h at 35°C, and 500 μl of *n*-hexane–diethyl ether (1:1 [vol/vol]) was added. Then, $Na_2S_2O_3$ (10% [m/vol]) was added to the mixture until the iodine was discolored. Subsequently, the organic phase obtained by centrifugation (8,000 $\times$ *g*, 5 min) was used for GC-MS analysis. One-microliter aliquots of the sample were analyzed with a similar GC-MS system by using the following temperature program: 50°C for 1 min, increased to 240°C at 15°C per min, held at 240°C for 20 min. The locations of the double bonds in the terminal olefins were deduced by comparing the fragmentation patterns of the control and DMDS derivatives.

Fatty acid feeding assay. Cyanobacterial strains were grown in 250-ml flasks containing 100 ml of BG11 medium with an initial optical density at 730 nm (OD_{730}) of ~0.1. Fatty acid stock solutions of pentadecanoic acid ($C_{15:0}$), heptadecanoic acid ($C_{17:0}$), or nonadecanoic acid ($C_{19:0}$) were prepared in dimethyl sulfoxide (DMSO) or ethanol and added at a final concentration of 0.1 mM. After 7 days of cultivation, the cells were harvested, and the hydrocarbons were analyzed.

Bioinformatic analyses. The deduced amino acid sequences of the Ols pathway from cyanobacteria were recovered from public databases. Multiple sequence alignments of the deduced amino acid sequences of the Ols pathway and CurM were performed using MAFFT (20) and visualized using UGENE software (21). Multiple sequence alignments of nucleotide and deduced amino acid sequences of Ols were conducted using Geneious, version 10.0.7 (22). Multiple sequence alignments of deduced amino acid sequences of FACLs and FAALs were analyzed using BioEdit, version 7.2.5 (23). The maximum likelihood trees were generated by MEGA 6.06 (Jones-Taylor-Thornton model, 1,000 bootstrap replicates) (24).

SUPPLEMENTAL MATERIAL

Supplemental material for this article may be found at https://doi.org/10.1128/AEM.00425-18.

SUPPLEMENTAL FILE 1, PDF file, 0.8 MB.

ACKNOWLEDGMENTS

This work was supported by the National Natural Science Foundation of China (grant 31570068 and 31200001 to T.Z.), the National Science Fund for Distinguished Young Scholars of China (grant 31525002 to X.L.), the Shandong Taishan Scholarship (to X.L.), and the Qingdao Innovative Leading Talent [15-10-3-15-(31)-zch to X.L.].

M.G. thanks the Institut Pasteur for funding the Pasteur Culture Collection of Cyanobacteria.

REFERENCES

1. Schirmer A, Rude MA, Li X, Popova E, del Cardayre SB. 2010. Microbial biosynthesis of alkanes. Science 329:559–562. https://doi.org/10.1126/science.1187936.
2. Liu AQ, Zhu T, Lu XF, Song LR. 2013. Hydrocarbon profiles and phylogenetic analyses of diversified cyanobacterial species. Appl Energy 111:383–393. https://doi.org/10.1016/j.apenergy.2013.05.008.
3. Li N, Chang WC, Warui DM, Booker SJ, Krebs C, Bollinger JM, Jr. 2012. Evidence for only oxygenative cleavage of aldehydes to alk(a/e)nes and formate by cyanobacterial aldehyde decarbonylases. Biochemistry 51:7908–7916. https://doi.org/10.1021/bi300912n.
4. Klahn S, Baumgartner D, Pfreundt U, Voigt K, Schon V, Steglich C, Hess WR. 2014. Alkane biosynthesis genes in cyanobacteria and their transcriptional organization. Front Bioeng Biotechnol 2:24. https://doi.org/10.3389/fbioe.2014.00024.
5. Coates RC, Podell S, Korobeynikov A, Lapidus A, Pevzner P, Sherman DH, Allen EE, Gerwick L, Gerwick WH. 2014. Characterization of cyanobacterial hydrocarbon composition and distribution of biosynthetic pathways. PLoS One 9:e85140. https://doi.org/10.1371/journal.pone.0085140.
6. Mendez-Perez D, Begemann MB, Pfleger BF. 2011. Modular synthase-encoding gene involved in alpha-olefin biosynthesis in *Synechococcus* sp. strain PCC 7002. Appl Environ Microbiol 77:4264–4267. https://doi.org/10.1128/AEM.00467-11.
7. Lea-Smith DJ, Ortiz-Suarez ML, Lenn T, Nurnberg DJ, Baers LL, Davey MP, Parolini L, Huber RG, Cotton CAR, Mastroianni G, Bombelli P, Ungerer P, Stevens TJ, Smith AG, Bond PJ, Mullineaux CW, Howe CJ. 2016. Hydrocarbons are essential for optimal cell size, division, and growth of cyanobacteria. Plant Physiol 172:1928–1940. https://doi.org/10.1104/pp.16.01205.
8. Kaczmarzyk D, Fulda M. 2010. Fatty acid activation in cyanobacteria mediated by acyl-acyl carrier protein synthetase enables fatty acid recycling. Plant Physiol 152:1598–1610. https://doi.org/10.1104/pp.109.148007.
9. Trivedi OA, Arora P, Sridharan V, Tickoo R, Mohanty D, Gokhale RS. 2004. Enzymic activation and transfer of fatty acids as acyl-adenylates in mycobacteria. Nature 428:441–445. https://doi.org/10.1038/nature02384.
10. Arora P, Goyal A, Natarajan VT, Rajakumara E, Verma P, Gupta R, Yousuf M, Trivedi OA, Mohanty D, Tyagi A, Sankaranarayanan R, Gokhale RS. 2009. Mechanistic and functional insights into fatty acid activation in *Mycobacterium tuberculosis*. Nat Chem Biol 5:166–173. https://doi.org/10.1038/nchembio.143.
11. Goyal A, Verma P, Anandhakrishnan M, Gokhale RS, Sankaranarayanan R. 2012. Molecular basis of the functional divergence of fatty acyl-AMP ligase biosynthetic enzymes of *Mycobacterium tuberculosis*. J Mol Biol 416:221–238. https://doi.org/10.1016/j.jmb.2011.12.031.
12. Liu Z, Ioerger TR, Wang F, Sacchettini JC. 2013. Structures of *Mycobacterium tuberculosis* FadD10 protein reveal a new type of adenylate-forming enzyme. J Biol Chem 288:18473–18483. https://doi.org/10.1074/jbc.M113.466912.
13. Gu LC, Wang B, Kulkarni A, Gehret JJ, Lloyd KR, Gerwick L, Gerwick WH, Wipf P, Hakansson K, Smith JL, Sherman DH. 2009. Polyketide decarboxylative chain termination preceded by *O*-sulfonation in curacin A biosynthesis. J Am Chem Soc 131:16033–16035. https://doi.org/10.1021/ja9071578.
14. Mendez-Perez D, Herman NA, Pfleger BF. 2014. A desaturase gene involved in the formation of 1,14-nonadecadiene in *Synechococcus* sp. strain PCC 7002. Appl Environ Microbiol 80:6073–6079. https://doi.org/10.1128/AEM.01615-14.
15. Rippka R, Deruelles J, Waterbury JB, Herdman M, Stanier RY. 1979. Generic assignments, strain histories and properties of pure cultures of *Cyanobacteria*. Microbiology 111:1–61. https://doi.org/10.1099/00221287-111-1-1.
16. Guan WN, Zhao H, Lu XF, Wang C, Yang ML, Bai FL. 2011. Quantitative analysis of fatty-acid-based biofuels produced by wild-type and genetically engineered cyanobacteria by gas chromatography-mass spectrometry. J Chromatogr A 1218:8289–8293. https://doi.org/10.1016/j.chroma.2011.09.043.
17. Tan XM, Yao L, Gao QQ, Wang WH, Qi FX, Lu XF. 2011. Photosynthesis driven conversion of carbon dioxide to fatty alcohols and hydrocarbons in cyanobacteria. Metab Eng 13:169–176. https://doi.org/10.1016/j.ymben.2011.01.001.
18. Lang I, Hodac L, Friedl T, Feussner I. 2011. Fatty acid profiles and their distribution patterns in microalgae: a comprehensive analysis of more than 2,000 strains from the SAG culture collection. BMC Plant Biol 11:124. https://doi.org/10.1186/1471-2229-11-124.
19. Vincenti M, Guglielmetti G, Cassani G, Tonini C. 1987. Determination of double-bond position in diunsaturated compounds by mass-spectrometry of dimethyl disulfide derivatives. Anal Chem 59:694–699. https://doi.org/10.1021/ac00132a003.
20. Katoh K, Kuma K, Toh H, Miyata T. 2005. MAFFT version 5: improvement in accuracy of multiple sequence alignment. Nucleic Acids Res 33:511–518. https://doi.org/10.1093/nar/gki198.
21. Okonechnikov K, Golosova O, Fursov M, UGENE Team. 2012. Unipro UGENE: a unified bioinformatics toolkit. Bioinformatics 28:1166–1167. https://doi.org/10.1093/bioinformatics/bts091.
22. Kearse M, Moir R, Wilson A, Stones-Havas S, Cheung M, Sturrock S, Buxton S, Cooper A, Markowitz S, Duran C, Thierer T, Ashton B, Meintjes P, Drummond A. 2012. Geneious Basic: an integrated and extendable desktop software platform for the organization and analysis of sequence data. Bioinformatics 28:1647–1649. https://doi.org/10.1093/bioinformatics/bts199.
23. Hall TA. 1999. BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. Nucleic Acids Symp Ser 41:95–98.
24. Tamura K, Stecher G, Peterson D, Filipski A, Kumar S. 2013. MEGA6: Molecular Evolutionary Genetics Analysis version 6.0. Mol Biol Evol 30:2725–2729. https://doi.org/10.1093/molbev/mst197.